PRIMARY AND SECONDARY CLEAVAGE SITES IN THE BAIT REGION OF α_2-MACROGLOBULIN

Steen B. MORTENSEN, Lars SOTTRUP-JENSEN*, H. Frede HANSEN, Torben E. PETERSEN and Staffan MAGNUSSON

Department of Molecular Biology, University of Aarhus, 8000 Århus C, Denmark

Received 26 October 1981

1. Introduction

The plasma glycoprotein α_2M forms complexes with proteinases from all 4 classes (serine, thiol, carboxyl and metalloproteinases, EC 3.4.21-24) [1–3]. Complex formation is initiated by specific limited proteolysis in the bait region, located near the middle of the 4 identical M_r 180 000 subunits [1–3]. Following cleavage in the bait region α_2M undergoes a change in conformation resulting in entrapment of the proteinase [2] and cleavage of the internal γ-glutamyl-β-cysteinyl thiol ester [4] in each subunit of α_2M [4–6]. A fraction of the proteinase molecules becomes covalently bound to α_2M [7,8]. The site of cleavage in the bait region of α_2M has been identified for trypsin [9,10], plasmin [9] and thrombin [9], respectively. In each case the same Arg–Leu bond was cleaved. Elastase was found to cleave at a Val–Gly bond (major cleavage) and at the adjacent Gly–Phe bond (minor cleavage) [9]. The new N-terminal sequence generated by cleavage of α_2M with *Staphylococcus aureus* strain V8 proteinase (SP) was also determined [10]. From these data and the sequence of the bait region [9] it is evident that cleavage in the bait region of α_2M by these different proteinases reflects their known substrate specificities [9,10].

Abbreviations: α_2M, α_2-macroglobulin (human); SP, *Staphylococcus aureus* strain V8 proteinase; TPCK, *N*-tosyl-L-phenylalanine–chloromethylketone; STI, soybean trypsin inhibitor (Kunitz); SGT, *Streptomyces griseus* trypsin-like proteinase; SGB, *Streptomyces griseus* proteinase B (chymotrypsin-like); Quadrol, *N,N,N',N'*-tetrakis-(2-hydroxypropyl) ethylene diamine; PTH, phenylthiohydantoin; HPLC, high-performance liquid chromatography

* To whom correspondence should be addressed

The common cleavage site for trypsin, plasmin, thrombin and the SP cleavage site (Glu–Ser) are located 15 and 5 residues, respectively, after the main elastase cleavage site [9,10]. Here, we report the identification of the cleavage sites in the bait region of α_2M for a variety of other proteinases with known primary structures from all 4 classes. We show that proteolytic cleavage can occur at each residue in the hexapeptide sequence: –Arg–Val–Gly–Phe–Tyr–Glu–. This sequence most likely contains the sites in the bait region at which initial proteolytic cleavage occurs. Additional probably secondary cleavage can occur in two other regions situated ~15 and 27 residues, respectively, from the hexapeptide region.

2. Materials and methods

Human α_2M was prepared by Zn^{2+}-affinity chromatography [4]. Bovine TPCK-treated trypsin, twice crystallized α-chymotrypsin and STI were from Worthington (Freehold NJ). Papain (crystal suspension, type III protease) and thermolysin (type X protease) were from Sigma (St Louis MO). *Streptomyces griseus* trypsin (SGT) was a gift from W. M. Award jr (Miami FL). *Streptomyces griseus* proteinase B (SGB) was a gift from L. Smillie (Edmonton, Canada). Crystalline subtilisin Novo (also known as subtilisin BPN') and subtilisin Carlsberg were gifts from I. Svendsen (Copenhagen). Calf chymosin (crystal suspension) was a gift from B. Foltmann (Copenhagen). Sephacryl S-300 was obtained from Pharmacia (Uppsala). Sequenator analysis of α_2M–proteinase complexes was performed as in [9] with a Beckman 890C instrument using 0.25 M Quadrol [11]. Low vacuum was

Published by Elsevier/North-Holland Biomedical Press
00145793/81/0000–0000/$02.75

applied for 140 s after cleavage with heptafluorobutyric acid (Beckman programme no. 122974). However, since the recovery of PTH-His and PTH-Arg was usually low under these conditions, that vacuum step was omitted in a number of Edman-degradations. Although this led to recovery of PTH-His and PTH-Arg in good yield, the repetitive yield was markedly decreased. For unexplained reasons the yield of PTH-Thr was very low in all degradations. PTH amino acids were analyzed by HPLC [12].

3. Experimental

Assuming ~60% active enzyme in the proteinase preparations used an amount of proteinase was added to 2 ml α_2M (7–10 mg/ml) which would give a 2-fold molar excess of active enzyme over α_2M. Incubations were performed at room temperature in 0.1 M Na-phosphate (pH 8.0) (except for α_2M-chymosin which was at pH 6.0) for the times shown in fig.2A,B. Then trichloroacetic acid was added to 15% (w/v). After 30 min at 0°C the precipitated α_2M–proteinase complexes were recovered by centrifugation and redissolved in 1–2 ml HCOOH [12].

The denaturated α_2M–proteinase complexes were then gel-chromatographed on Sephacryl S-300 using 50% HCOOH as eluent. The material in the irregular void volume peak [9] was diluted with 10 vol. water and freeze-dried. N-Terminal sequence determination of α_2M–proteinase complexes (4–8 mg) could be carried out through only 10–12 steps due to the rapidly increasing 'background' and 'carry over' [9]. The yields of the PTH-amino acids recovered in each step were plotted as shown in fig.1. In most cases this permitted discrimination of even minor new N-terminal sequences from 'background'.

4. Results and discussion

Fig.1 shows that a second new N-terminal sequence: Val–Gly–Phe–Tyr–X–X–Asp– is present in the α_2M–trypsin complex formed in the presence of 20 mM benzamidine in addition to the new N-terminal sequence: Leu–Val–His–Val–Glu–Glu– and the original N-terminal sequence of α_2M: Ser–Val–Ser–Gly–Lys–Pro– [9]. Fig.2 summarizes the interpretation of the results of sequence determination of the α_2M–proteinase complexes studied here, utilizing the

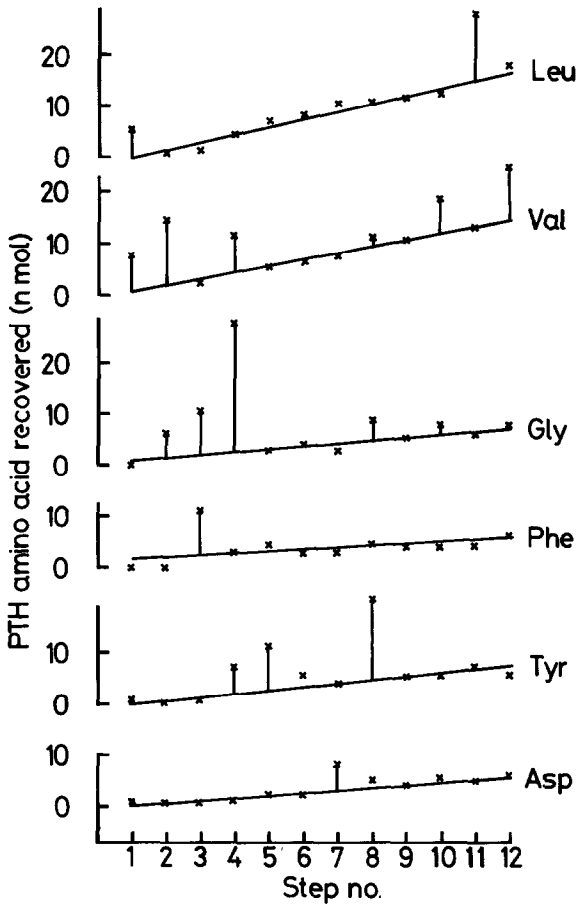

Fig.1. Yield of PTH-Leu, -Val, -Gly, -Phe, -Tyr and -Asp, respectively, in each step of Edman degradation of α_2M–trypsin complex formed in the presence of 20 mM benzamidine. The data show the steadily increasing background, most pronounced for PTH-Leu and PTH-Val. The yield of the PTH amino acids in each step was calculated by deducting the interpolated background value (——) from the total yield.

known amino acid sequence of the bait region [9]. Fig.3 gives the sequence of the bait region (residues 1–35 from [9]) and the following 10 residues. For trypsin it is evident that the Arg_{22}–Leu bond identified earlier as the site of cleavage in the bait region is not the only site since cleavage was also observed at Arg_7–Val when trypsin was strongly competitively inhibited by benzamidine. This additional site of cleavage was also identified in α_2M–trypsin complex formed at pH 6.0, although in lower yield (not shown). Furthermore, the results of sequence determination on α_2M–SGT complex formed in the presence of benzamidine show that cleavage has again

A

α_2M-trypsin (20 mM benzamidine, 2 min)

1: Ser-Val-Ser-Gly-Lys-Pro-Gln-Tyr-Met-Val-Leu-Val-
(1) (13) (5) (15)(19)(22)(13)(16) (7) (7) (13)(11)

2: Leu-Val-His-Val-Glu-Glu-Pro-His-Thr-Glu-Thr-Val-
(6) (-) (7) (13) (9) (11) (-) (-) (8) (-)

3: Val-Gly-Phe-Tyr-Glu-Ser-Asp-Val-Met-Gly-Arg-Gly-
(8) (5) (9) (5) (1) (6) (2) (2) (-) (1)

α_2M-SGT (20 mM benzamidine, 2 min)

1: Ser-Val-Ser-Gly-Lys-Pro-Gln-Tyr-Met-Val-
(1) (3) (1) (4) (4) (4) (4) (3) (3) (1)

2: Leu-Val-His-Val-Glu-Glu-Pro-His-Thr-Glu-
(1) (1) (1) (6) (2) (2) (1) (-) (2)

3: Val-Gly-Phe-Tyr-Glu-Ser-Asp-Val-Met-Gly
(2) (1) (2) (3) (-) (3) (1) (4)

α_2M-chymotrypsin (STI:CT = 4 mol/mol, 3 min)

1: Ser-Val-Ser-Gly-Lys-Pro-Gln-Tyr-Met-Val-
(4) (23) (7) (28)(11)(15)(14)(17)(14)(13)

2: Glu-Ser-Asp-Val-Met-Gly-Arg-Gly-His-Ala-
(7) (2) (8) (8) (10) (7) (-) (5) (-) (7)

α_2M-SGB (STI:SGB = 4 mol/mol, 1 min)

1: Ser-Val-Ser-Gly-Lys-Pro-Gln-Tyr-Met-Val-
(1) (6) (1) (11) (9) (16) (8) (6) (5) (3)

2: Val-His-Val-Glu-Glu-Pro-His-Thr-Glu-Thr-
(5) (8) (5) (9) (10) (6) (-) (3) (1)

α_2M-subtilisin Novo (1 min)

1: Ser-Val-Ser-Gly-Lys-Pro-Gln-Tyr-Met-Val-Leu-Val-
(1) (11) (6) (6) (7) (10) (8) (6) (3) (5) (4) (4)

2: Leu-Val-His-Val-Glu-Glu-Pro-His-Thr-Glu-Thr-Val-
(3) (5) (6) (6) (3) (5) (5) (1) (2) (1)

3: Val-His-Val-Glu-Glu-Pro-His-Thr-Glu-Thr-Val-Arg-
(4) (4) (4) (3) (4) (-) (1) (-) (1) (-)

B

α_2M-papain (1 min)

1: Ser-Val-Ser-Gly-Lys-Pro-Gln-Tyr-Met-Val-Leu-
(1) (14) (3) (11) (4) (7) (5) (7) (10) (6) (9)

2: Phe-Tyr-Glu-Ser-Asp-Val-Met-Gly-Arg-Gly-His-
(2) (2) (1) (-) (1) (1) (2) (2) (-) (2) (-)

3: Val-Gly-Phe-Tyr-Glu-Ser-Asp-Val-Met-Gly-Arg-
(2) (2) (1) (2) (1) (-) (-) (-) (-)

α_2M-papain (5 min)

1: Ser-Val-Ser-Gly-Lys-Pro-Gln-Tyr-Met-Val-Leu-
(1) (6) (1) (6) (7) (7) (3) (4) (6) (4) (5)

2: Val-Glu-Glu-Pro-His-Thr-Glu-Thr-Val-Arg-Lys-
(6) (7) (8) (8) (8) (-) (6) (-) (6) (5) (5)

α_2M-chymosin (1 min)

1: Ser-Val-Ser-Gly-Lys-Pro-Gln-Tyr-Met-Val-Leu-Val-
(4) (14) (9) (18)(16)(14)(12)(10)(11) (5) (11) (9)

2: Tyr-Glu-Ser-Asp-Val-Met-Gly-Arg-Gly-His-Ala-Arg-
(12)(13) (7) (10) (5) (4) (3) (8) (-) (8) (2)

3: Arg-Lys-Tyr-Phe-Pro-Glu-Thr-Trp-Ile-Trp-Asp-Leu-
(1) (7) (4) (3) (5) (3) (1) (-) (-) (-) (1) (5)

4: Phe-Pro-Glu-
(2) (1) (1)

α_2M-thermolysin (1 min)

1: Ser-Val-Ser-Gly-Lys-Pro-Gln-Tyr-Met-Val-Leu-Val-
(2) (30) (5) (19)(16)(12) (8) (11)(12) (9) (14)(17)

2: Leu-Val-His-Val-Glu-Glu-Pro-His-Thr-Glu-Thr-Val-
(21) (7) (13)(16)(13) (9) (1) (1) (10) (1)

Fig.2. Summary of the interpretation of the sequence results for α_2M complexed with proteinases: (A) the serine proteinases trypsin, SGT, chymotrypsin, SGB and subtilisin Novo, respectively; (B) the thiol proteinase papain, the carboxyl proteinase chymosin and the metallo proteinase thermolysin, respectively; (1) the original N-terminal sequence of α_2M [9,10,14]; (2) the new N-terminal sequence following complex formation; (3,4) the additional N-terminal sequences observed in some cases. The values listed in brackets denote the yield of PTH-amino acid recovered. Lack of quantitative data is indicated as (–). The time of incubating α_2M with proteinase before addition of trichloroacetic acid is indicated.

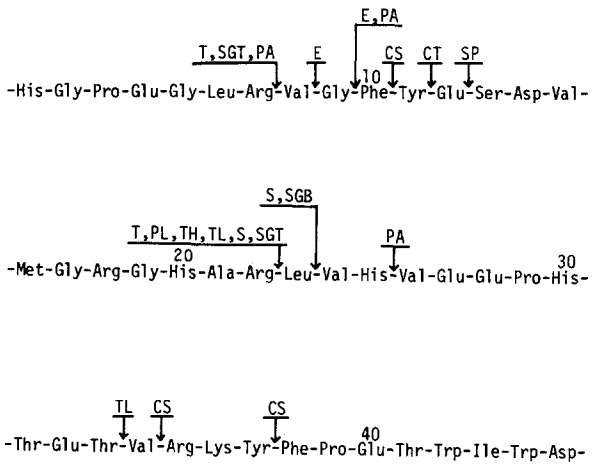

Fig.3. Amino acid sequence of a 45 residue stretch of α_2M containing the bait region. The sequence of the 10 residues that follow after the reported sequence [9] was deduced from the sequence of the peptide Thr–Val–Arg–Lys–Tyr–Phe–Pro–Glu, obtained from CB-P [9] after cleavage with *Staphylococcus aureus* (SP) proteinase and the sequence of a large tryptic peptide having the N-terminal sequence: Lys–Tyr–Phe–Pro–Glu–Thr–Trp–Ile–Trp–Asp– (L. S., unpublished). The cleavage sites for different proteinases are shown by vertical arrows: T, bovine trypsin; SGT, *Streptomyces griseus* trypsin; PA, papain; E, porcine elastase; CS, calf chymosin; CT, bovine chymotrypsin; SP, *Staphylococcus aureus* strain V8 proteinase (data from [10]); PL, human plasmin; TH, bovine thrombin; TL, thermolysin; S, subtilisin Novo; SGB, *Streptomyces griseus* proteinase B.

occurred at the same residues (Arg_{22}–Leu and Arg_7–Val). These results show that the cleavage at Arg_{22}–Leu observed earlier is likely to be the result of secondary proteolytic cleavage in the bait region following the primary cleavage at Arg_7–Val.

The cleavage sites for chymotrypsin could not be deduced from the sequence on the α_2M–chymotrypsin complex, even when this complex had been formed in the presence of the competitive inhibitor *N*-acetyl-L-tryptophan, since a large number of N-termini were generated. However, by forming the α_2M–chymotrypsin complex in the presence of STI which is also a moderately strong competitive inhibitor for chymotrypsin [13], a single site of cleavage in the bait region of α_2M was identified under these conditions: Tyr_{11}–Glu.

For α_2M–SGB complex formed in the presence of STI a single site of cleavage in the bait region of α_2M was identified: Leu_{23}–Val.

In the case of α_2M–subtilisin Novo complex 2 sites of cleavage in the bait region of α_2M were identified: Arg_{22}–Leu and Leu_{23}–Val. Subtilisin Carlsberg also cleaved α_2M at these sites (not shown). The experiments with the latter 2 proteinases were performed without inhibitors. Since any small peptides released would have been separated from the α_2M–proteinase complex in the trichloroacetic acid precipitation or gel-chromatography steps it is not clear whether the cleavage sites identified here are primary or secondary.

Two separate experiments were performed with the α_2M–papain complex, using a complex formation period of 1 min and 5 min, respectively. In the 1 min complex cleavage at Arg_7–Val and Gly_9–Phe in the bait region of α_2M was obtained in rather low yield. However, in the 5 min complex cleavage at these 2 sites could not be detected. Only the new N-terminal sequence due to cleavage of His_{25}–Val was observed. This shows that the primary cleavages at Arg_7–Val and Gly_9–Phe are followed by a rapid secondary cleavage at His_{25}–Val.

For α_2M–chymosin complex a major site of cleavage in the bait region of α_2M was found: Phe_{10}–Tyr. Two minor sites of cleavage could also be identified: Val_{34}–Arg and Tyr_{37}–Phe.

In one experiment with α_2M–thermolysin complex, cleavage had occurred at the Arg_{22}–Leu bond in the bait region of α_2M. In another experiment an additional minor cleavage at Thr_{33}–Val was also observed (not shown).

In addition to the N-terminal sequences originating from α_2M minor N-terminal sequences originating from the proteinases used could also be discriminated. The residues that could be clearly attributed to each proteinase by comparing the minor N-terminal sequences of the α_2M–proteinase complexes with those of the respective proteinases are shown in fig.4. The yield was 0.25–0.35 mol/mol α_2M subunit for trypsin and SGT, while it was only 0.10–0.20 mol/mol α_2M subunit for the other proteinases. However, appearance of 3.0–4.0 mol SH/mol α_2M [4] was observed with all enzymes used.

This work adds further evidence to the results in [9,10] showing that the site(s) of bait region cleavage in α_2M reflects the known substrate specificity of the particular proteinase used. So far unusual cleavages have been only observed for subtilisin and chymosin (cleavage at Arg_{22}–Leu and Val_{34}–Arg, respectively).

As seen in fig.3 the sites of cleavage found so far are located in 3 discrete parts of the bait region, namely:

T: Ile-Val-Gly-Gly-Tyr-Thr-Cys-Gly-Ala-Asn-

CT: Ile-Val-Asn-Gly-Glu-Glu-Ala-Val-Pro-Gly-Ala-Asn-Thr-Pro-Asp-Arg-Leu-Gln-Gln-Ala

SGT: Val-Val-Gly-Gly-Thr-Arg-Ala-Ala-Gln-Gly-

SGB: Ile-Ser-Gly-Gly-Asp-Ala-Ile-Tyr-Ser-Ser-

S(Novo): Ala-Gln-Ser-Val-Pro-Tyr-Gly-Val-Ser-Gln-

PA: Ile-Pro-Glu-Tyr-Val-Asp-Trp-Arg-Gln-Lys-

CS: Gly-Glu-Val-Ala-Ser-Val-Pro-Leu-Thr-Asn-

TL: Ile-Thr-Gly-Thr-Ser-Thr-Val-Gly-Val-Gly

Fig.4. Identification of minor sequences in α_2M–proteinase complexes that can be attributed to the particular proteinase. The sequences of the first 10 residues from each proteinase are from [15–22]. Those residues which are underlined could be clearly identified because they occur only in the proteinase sequences in a particular step. For chymotrypsin only the sequences of the B and C chains could be followed while that of the A-chain was not identified.

(1) At each residue of the hexapeptide sequence –Arg_7–Val–Gly–Phe–Tyr–Glu– (trypsin, SGT, papain, elastase, chymosin, chymotrypsin and SP [10];
(2) At Arg_{22}–Leu (trypsin, plasmin, thrombin, thermolysin, subtilisin Novo and Carlsberg and SGT), Leu_{23}–Val (subtilisin Novo and Carlsberg and SGB and His_{25}–Val (papain);
(3) At Thr_{33}–Val (thermolysin), Val_{34}–Arg and Tyr_{37}–Phe (chymosin).

The finding that short incubation of α_2M with papain results in cleavage of the Arg_7–Val and Gly_9–Phe bonds while a longer incubation exposes the sequence Val_{26}–Glu–Glu– shows that the primary bait region cleavage by papain occurs in the hexapeptide sequence. Furthermore, cleavage by trypsin or SGT at Arg_7–Val can be detected under conditions (e.g., in 20 mM benzamidine or at pH 6.0) where the activity of these proteinases is strongly suppressed. Finally, only the Tyr_{11}–Glu bond was found to be cleaved, when chymotrypsin was inhibited by STI. Thus, this hexapeptide part of the bait region contains specific sites for primary cleavage by a large number of proteinases. For those enzymes where cleavage has only been observed in the 2 other parts of the bait region (~15 and 27 residues, respectively, from the main elastase cleavage site) it is still conceivable that the primary cleavage has actually occurred in the same hexapeptide stretch, the sequence of which fits the known specificity requirements of all these proteinases.

As summarized in [2], α_2M forms complexes most readily with proteinases that have a broad substrate specificity. However, mammalian collagenases, which cleave the α_1 (I) chain of collagen at only one site [23] can also cleave α_2M near the middle of its subunits and form complexes with α_2M [3]. The similarity between the sequence of the cleavage site in collagen: –Gly–Pro–Gln–Gly–Ile– [23,24] and the sequence –Gly–Pro–Glu–Gly–Leu adjacent to the hexapeptide sequence in the bait region of α_2M suggests that the site of cleavage by this collagenase in α_2M is probably the Gly_5–Leu bond.

The activation of α_2M by specific limited proteolysis in the bait region leads to the formation of an intermediate form ('nascent' α_2M) in which the Glx-residues of the activated thiol esters can form covalent linkages to the bound proteinase [4,8]. In this respect, the activation of α_2M is reminiscent of other processes regulated by specific limited proteolysis, e.g., zymogen activation and fibrinogen conversion. In contrast with the very restricted cleavage specificities found in these systems the clustering of a number of determinants for specific cleavage in the bait region allows α_2M to be activated by a variety of proteinases with different catalytic sites and mechanisms. The resulting α_2M complexes probably serve to remove the particular proteinase from the circulation [25].

Acknowledgements

We thank Drs W. M. Award jr, L. Smillie, I. Svendsen and B. Foltmann for kind gifts of highly purified proteinases. This work was supported by The National Heart, Lung and Blood Institute, NIH (Bethesda MD) grant no. HL 16238.

References

[1] Harpel, P. C. (1973) J. Exp. Med. 138, 508–521.
[2] Barrett, A. J. and Starkey, P. M. (1973) Biochem. J. 133, 709–724.
[3] Werb, Z., Burleigh, M. C., Barrett, A. J. and Starkey, P. M. (1974) Biochem. J. 139, 359–368.
[4] Sottrup-Jensen, L., Petersen, T. E. and Magnusson, S. (1980) FEBS Lett. 121, 275–279.
[5] Salvesen, G. S., Sayers, C. A. and Barrett, A. J. (1981) Biochem. J. 195, 453–461.

[6] Howard, J. B. (1981) Proc. Natl. Acad. Sci. USA 78, 2235–2239.
[7] Salvesen, G. S. and Barrett, A. J. (1980) Biochem. J. 187, 695–701.
[8] Sottrup-Jensen, L., Petersen, T. E. and Magnusson, S. (1981) FEBS Lett. 128, 127–132.
[9] Sottrup-Jensen, L., Lønblad, P. B., Stepanik, T. M., Petersen, T. E., Magnusson, S. and Jörnvall, H. (1981) FEBS Lett. 127, 167–173.
[10] Hall, P. K., Nelles, L. P., Travis, J. and Roberts, R. C. (1981) Biochem. Biophys. Res. Commun. 100, 8–16.
[11] Sottrup-Jensen, L., Hansen, H. F., Mortensen, S. B., Petersen, T. E. and Magnusson, S. (1981) FEBS Lett. 123, 145–148.
[12] Sottrup-Jensen, L., Petersen, T. E. and Magnusson, S. (1980) Anal. Biochem. 107, 456–460.
[13] Wu, F. C. and Laskowski, M. sr (1955) J. Biol. Chem. 213, 609–614.
[14] Swenson, R. P. and Howard, J. B. (1979) J. Biol. Chem. 254, 4452–4456.
[15] Walsh, K. A., Kauffman, D. L., Sampath Kumar, K. S. V. and Neurath, H. (1964) Proc. Natl. Acad. Sci. USA 51, 301–304.
[16] Hartley, B. S. and Kauffman, D. L. (1966) Biochem. J. 101, 229–240.
[17] Olafson, R. W., Jurasek, L., Carpenter, M. R. and Smillie, L. B. (1975) Biochemistry 14, 1168–1177.
[18] Jurasek, L., Carpenter, M. R., Smillie, L. B., Gertler, A., Levey, S. and Ericsson, L. H. (1974) Biochem. Biophys. Res. Commun. 61, 1095–1100.
[19] Markland, F. S. jr and Smith, E. L. (1971) in: The Enzymes (Boyer, P. D. ed) vol. 13, pp. 561–608, Academic Press, New York.
[20] Mitchell, R., Chaiken, I. M. and Smith, E. L. (1970) J. Biol. Chem. 245, 3485–3492.
[21] Foltmann, B., Petersen, V. B., Kauffman, D. and Wybrandt, G. (1979) J. Biol. Chem. 254, 8447–8456.
[22] Titani, K., Hermodson, M. H., Ericsson, L. H., Walsh, K. A. and Neurath, H. (1972) Nature 238, 35–37.
[23] Gross, J., Harper, E., Harris, E. D., McCroskery, P. A., Highberger, J. H., Corbett, C. and Kang, A. H. (1974) Biochem. Biophys. Res. Commun. 61, 605–612.
[24] Fietzek, P. P., Rexrodt, F. W., Hopper, K. E. and Kühn, K. (1973) Eur. J. Biochem. 38, 396–400.
[25] Ohlsson, K. (1976) Prot. Biol. Fluids 23, 43–45.